Personality and Social Psychology Bulletin

http://psp.sagepub.com/

Accounting for Successful Control of Implicit Racial Bias : The Roles of Association Activation, Response Monitoring, and Overcoming Bias

Karen Gonsalkorale, Jeffrey W. Sherman, Thomas J. Allen, Karl Christoph Klauer and David M. Amodio

Pers Soc Psychol Bull 2011 37: 1534 originally published online 21 July 2011
DOI: 10.1177/0146167211414064

The online version of this article can be found at:
http://psp.sagepub.com/content/37/11/1534

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Society for Personality and Social Psychology

Article

Personality and Social Psychology Bulletin
37(11) 1534–1545

DOI: 10.1177/0146167211414064
http://pspb.sagepub.com
SAGE

Accounting for Successful Control of Implicit Racial Bias: The Roles of Association Activation, Response Monitoring, and Overcoming Bias

Karen Gonsalkorale[1], Jeffrey W. Sherman[2], Thomas J. Allen[2], Karl Christoph Klauer[3], and David M. Amodio[4]

Abstract

Individuals who are primarily internally motivated to respond without prejudice show less bias on implicit measures than individuals who are externally motivated or unmotivated to respond without prejudice. However, it is not clear why these individuals exhibit less implicit bias than others. We used the Quad model to examine motivation-based individual differences in three processes that have been proposed to account for this effect: activation of associations, overcoming associations, and response monitoring. Participants completed an implicit measure of stereotyping (Study 1) or racial attitudes (Study 2). Modeling of the data revealed that individuals who were internally (but not externally) motivated to respond without prejudice showed enhanced detection and reduced activation of biased associations, suggesting that these processes may be key to achieving unbiased responding.

Keywords

implicit prejudice, motivation to respond without prejudice, implicit attitudes, prejudice, implicit bias

Received September 12, 2010; revision accepted May 11, 2011

Prejudices and stereotypes are often conceived as ubiquitous, but people differ in the extent to which they show biased attitudes and beliefs. One important source of individual differences in bias is motives and goals relating to the control and expression of prejudice (e.g., Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Dunton & Fazio, 1997; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999; Moskowitz, Salomon, & Taylor, 2000). People who have chronic egalitarian goals (Moskowitz et al., 1999; Moskowitz et al., 2000) or are motivated to control prejudice (Dunton & Fazio, 1997) respond with less bias on self-reported and implicit measures of stereotyping and prejudice than people who do not have egalitarian goals or are low in motivation to control prejudice. Other research suggests that biased responding depends not on an individual's level of motivation to respond without prejudice but on the reasons underlying these motives. People can be motivated to respond without prejudice primarily for internal reasons (i.e., being egalitarian is important to the self-concept), external reasons (i.e., to avoid social disapproval), or both internal and external reasons (Plant & Devine, 1998). Individuals who are either internally or externally motivated to respond without prejudice show lower levels of explicit bias than individuals who are not motivated to respond without prejudice (Plant & Devine, 1998). However, only those individuals who are internally but not externally motivated are able to respond without bias on implicit measures (Amodio, Devine, & Harmon-Jones, 2008; Amodio, Harmon-Jones, & Devine, 2003; Devine et al., 2002).

Although the importance of motives and goals in biased responding is widely acknowledged, questions remain as to why some people are less prone than others to implicit bias. The current research examines this issue in the context of motivation-based individual differences in implicit stereotyping and prejudice. Specifically, the central goal of the present research is to better understand exactly why those with high internal but low external motivations show less bias on implicit measures. The question at the heart of this

[1]University of Sydney, Sydney, Australia
[2]University of California, Davis, Davis, CA, USA
[3]University of Freiburg, Freiburg, Germany
[4]New York University, New York, NY, USA

Corresponding Author:
Karen Gonsalkorale, University of Sydney, School of Psychology, Brennan MacCallum Building, NSW, Australia, 2006
Email: karen.gonsalkorale@sydney.edu.au

issue is whether individuals differ in the extent to which they have race-biased associations that are activated automatically. On one hand, individuals who are motivated for internal (but not external) reasons may simply not have biased associations activated to the same extent as other individuals. On the other hand, it may be that biased associations are activated to the same degree among all members of a society, but that these associations are more likely to translate into biased responses for some people than others. That is, individuals may differ in the extent to which they are able to control their responses, either by overcoming the effects of biased associations or by monitoring their behavior to reduce biased responding. Of course, it also is possible that individual differences in biased responding emerge from some combination of these factors. As outlined below, elements of these different accounts of motivation-based differences in implicit bias have been featured in the key explanations of these findings.

Explaining Motivation-Based Individual Differences in Implicit Bias

Automatic Activation of Associations

One possibility is that motivation-based differences in implicit bias reflect differences in the extent to which individuals possess biased associations that are automatically activated. This idea has been used to explain the finding that participants who are internally (but not externally) motivated to respond without prejudice do not show bias on implicit measures of racial attitudes (Devine et al., 2002). Similarly, the finding that these individuals do not exhibit bias on a physiological measure of affective race bias (indexed via startle eyeblink response to Black and White faces) has been interpreted as evidence that some individuals do not have biased affective associations activated in the first place (Amodio et al., 2003).

It may be that individuals who have chronic egalitarian goals or are internally motivated to respond without prejudice were simply less likely to have acquired biased associations in the first place (e.g., during childhood or later development). Alternatively, several researchers have proposed that appropriately motivated individuals fail to exhibit automatically activated biased associations because such biases are automatically inhibited (Glaser & Knowles, 2008; Johns, Cullum, Smith, & Freng, 2008; Moskowitz et al., 1999; Moskowitz et al., 2000; Park, Glaser, & Knowles, 2008). The argument is that prejudice-relevant stimuli (e.g., Black faces) automatically activate egalitarian goals for these individuals, which help them to inhibit activation of biased associations (e.g., Johns et al., 2008). The fact that these individuals show less bias on implicit measures has been taken as evidence that otherwise automatically activated associations must have been inhibited (automatically). Irrespective of the proposed underlying mechanism, explanations focusing on automatic activation of associations share the common assumption that biased associations are less activated among individuals who have chronic egalitarian goals or are internally motivated to respond without prejudice.

Response Monitoring

Other researchers have proposed a different account of how motivations to respond without prejudice influence responses on implicit measures. Monteith, Ashburn-Nardo, Voils, and Czopp's (2002) model of prejudice reduction proposes that a key difference between people who can and cannot control bias lies in their ability to monitor responses to make sure that they are in line with egalitarian goals. Although this theory was originally formulated with regard to individual differences in explicit prejudice, response monitoring should also be important on implicit measures. Support for this idea was provided by recent work that investigated conflict monitoring (which is one important kind of response monitoring) among individuals who differed in their motivation to respond without prejudice. Amodio et al. (2004) proposed that effective control over implicit race bias is dependent on two neurocognitive systems: the conflict monitoring system (e.g., Carter et al., 1998) and the regulatory system (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001). The conflict monitoring system functions to detect instances of competing responses. When conflict is detected, this system is thought to alert the regulatory system, which functions to strengthen the intended (i.e., nonprejudiced) response, thereby overriding the unintended response. Amodio et al. (2008, Study 1) explored the role of conflict monitoring by measuring event-related potentials (ERPs) as participants performed the Weapons Identification Task (Payne, 2001), an implicit measure of stereotyping. Participants' motivations to respond without prejudice had been measured in a previous session using Plant and Devine's (1998) internal motivation scale (IMS) and external motivation scale (EMS). Analyses focused on a neural index of conflict monitoring called the error-related negativity (ERN), a component of the ERP that has been linked to activity in the anterior cingulate cortex, which in turn has been associated with conflict monitoring processes in a large body of research (for a review, see Botvinick et al., 2001). ERN amplitudes are most pronounced during response errors on the task, as errors represent the most extreme conflict between an intended response and an unfolding unintended behavior. Amodio et al.'s results showed that, when participants made stereotypical errors (i.e., misidentifying a tool as a gun following a Black face), high IMS–low EMS participants demonstrated enhanced ERN amplitudes compared with high IMS–high EMS participants and low IMS participants. Hence, high IMS–low EMS participants showed stronger conflict monitoring when their responses were discrepant with the goal to be nonprejudiced, suggesting that these individuals may be successful at responding without prejudice because they are more adept at detecting competition between appropriate and inappropriate responses.

Amodio et al. (2008) used the Process Dissociation procedure (PD; Jacoby, 1991; Payne, 2001) to provide further support for their conclusion. PD is a behavior modeling technique that extracts independent estimates of automatic bias and controlled processing from performance on implicit measures. The controlled process estimated by PD reflects the outcome of control, such that it represents a person's success in responding accurately in line with intentions. Amodio et al. (2008; also see Schlauch, Lang, Plant, Christensen, & Donohue, 2009) reported higher PD-control estimates among high IMS–low EMS participants than high IMS–high EMS participants, who, in turn, showed higher levels than low IMS participants. In addition, research by Amodio and colleagues (Amodio et al., 2004; Amodio et al., 2008; Amodio, Kubota, Harmon-Jones, & Devine, 2006) has shown that stronger conflict monitoring responses, as measured by the ERN, are associated with greater PD estimates of controlled behavior and that IMS–EMS differences on the PD estimate of control corresponded with difference in ERN amplitudes. These PD analyses showed no differences in automatic bias (the other process estimated by PD) across levels of motivations.

Inhibition

A third possibility is that individuals differ in the extent to which they are able to inhibit automatic biases. As described above, a number of researchers have argued that such inhibition processes are responsible for individual differences in the prevalence of biased associations (and, therefore, in the extent of implicit bias; Glaser & Knowles, 2008; Johns et al., 2008; Moskowitz et al., 1999; Moskowitz et al., 2000; Park et al., 2008). Note, however, that it also is possible to reduce biased behavioral responses via successful inhibition without necessarily reducing the activation of associations. That is, inhibition may influence behavior independently of its effects on association activation. Monteith (e.g., Monteith et al., 2002) has argued that behavioral inhibition is a central component of preventing biased responses and that, for nonbiased responding to occur, individuals must not only recognize when control is needed (via response monitoring) but also inhibit biased behaviors that are triggered by activated associations. This suggests that chronic egalitarians and internally (but not externally) motivated individuals may be better able to inhibit the effects of biased associations than other individuals.

Consistent with these hypotheses, research has shown that individuals with an internalized egalitarian goal were less cognitively depleted after suppressing stereotype use than individuals who did not have an egalitarian goal, suggesting efficient inhibitory skills (Gordijn, Hindriks, Koomen, Dijksterhuis, & van Knippenberg, 2004). In addition, whereas cognitive load manipulations lead to stronger patterns of bias on implicit measures of stereotyping and prejudice for individuals who are low in implicit motivation to control prejudice (Park et al., 2008) or self-determination to respond without prejudice (Legault, Green-Demers, & Eadie, 2009), they have no effect on individuals who are high on internal but low on external motivations to respond without prejudice (Devine et al., 2002, Study 3). To the extent that load interferes with attempts to inhibit biased behavior, as proposed by these researchers, these results suggest that properly motivated individuals are more efficient inhibitors. Finally, another study found that when individuals were intoxicated (thereby reducing their cognitive capacity), those who were internally but not externally motivated to respond without prejudice made fewer race-biased errors on a measure of implicit stereotyping than other individuals and also demonstrated greater control over responses than other individuals (Schlauch et al., 2009).

Unanswered Questions

Key aspects of the explanations for motivation-based differences in implicit bias have not been tested directly, leaving many questions unanswered. The idea that biased associations are successfully inhibited by some individuals (e.g., Glaser & Knowles, 2008; Moskowitz et al., 1999; Moskowitz et al., 2000; Park et al., 2008) is based primarily on evidence of less implicit bias among these individuals. That is, less implicit bias is assumed to reflect the inhibition of associations. However, although reduced implicit bias may reflect the successful inhibition of associations, it may also reflect the possibility that bias is simply not activated in the first place (or activated in an opposing direction), independent of any inhibitory processes. Alternatively, reduced bias also may reflect differences in the extent to which individuals are able to inhibit the influence of racial associations on their behavior when completing implicit measures.

It is difficult to determine the precise mechanisms underlying a nonprejudiced response on implicit measures because such measures do not tap the activation or inhibition of associations alone. A growing body of research has shown that performance on implicit measures reflects multiple component processes, including monitoring for appropriate responses (Allen, Sherman, & Klauer, 2010; Amodio et al., 2004; Bartholow, Dickter, & Sestir, 2006; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Gonsalkorale, Allen, Sherman, & Klauer, 2010; Gonsalkorale, Sherman, & Klauer, 2009; Gonsalkorale, von Hippel, Sherman, & Klauer, 2009; Payne, 2001; Sherman, 2006; Sherman et al., 2008). Thus, scores on implicit measures may reflect activation of associations, the ability to overcome those associations, the ability to effectively monitor responses when completing the measure, or some combination of these processes. This means that motivation-based individual differences in implicit bias may reflect any combination of these processes. To date, no research has independently examined these three theoretical accounts of the relationship between motives to respond without prejudice and the expression of implicit bias. This was the goal of the present research.

The Quad Model

One approach to examining the separate processes reflecting the activation of associations, the ability to overcome them, and response monitoring is suggested by the Quadruple Process model (Quad model; Conrey et al., 2005; Sherman et al., 2008). The Quad model is a multinomial model (see Batchelder & Riefer, 1999) designed to estimate the independent contributions of multiple processes from responses on implicit measures of bias (for reviews of this approach, see Sherman, 2006; Sherman et al., 2008). According to the model, responses on implicit measures of bias reflect the operation of four qualitatively distinct processes: Activation of Associations (AC), Detection of Responses (D), Overcoming Bias (OB), and Guessing (G). The AC parameter refers to the degree to which biased associations are automatically activated when responding to a stimulus. All else being equal, the stronger the associations, the more likely they are to be activated and influence behavior. The D parameter reflects a relatively controlled process that discriminates between appropriate and inappropriate responses. Sometimes, the activated associations conflict with the detected correct response. For example, on incompatible trials of implicit stereotyping measures (e.g., a Black face followed by a tool in the Weapons Identification Task; Payne, 2001), automatically activated racial associations (e.g., between Blacks and guns) would conflict with the detected correct response (i.e., “tool”). In such cases, the Quad model proposes that an OB process resolves the conflict. As such, the OB parameter refers to inhibitory processes that prevent automatically activated associations from influencing behavior when they conflict with detected correct responses. Finally, the G parameter reflects general response tendencies that may occur when individuals have no associations that direct behavior and they are unable to detect the appropriate response. The Quad model and the construct validity of its parameters have been extensively validated in previous research (see Beer et al., 2008; Conrey et al., 2005; Gonsalkorale, Sherman, et al., 2009; Sherman et al., 2008).

Overview of the Present Research

In two studies, we applied the Quad model to examine motivation-based individual differences in implicit stereotyping and prejudice. More specifically, we used the Quad model to test three competing theoretical accounts of why individuals who are internally (but not externally) motivated to respond without prejudice exhibit less implicit bias than others. The three accounts postulate that people with high IMS–low EMS differ from others in terms of activation of automatic associations, inhibition of biased associations, or monitoring of appropriate and inappropriate responses. Of course, it also is possible that a combination of these factors is important. If individual differences in implicit bias are a function of automatic associations, process estimates of AC should be lower for high IMS–low EMS individuals than for other individuals. If high IMS–low EMS individuals are proficient at inhibiting automatic associations, then these individuals should show higher levels of OB. However, there should instead be individual differences in D if high IMS–low EMS individuals are especially able to monitor their responses. In addition to the primary aim of testing these alternative accounts, we also were able to examine motivation-based differences in G. Thus, our modeling approach allowed us to examine for the first time the roles of four central processes in motivation-based differences in implicit bias. A significant advantage of this approach is that we were able to assess the underlying processes simultaneously and independently in a single task.

Study 1

In Study 1 we reanalyzed data from Amodio et al. (2008, Study 1) using the Quad model to examine individual differences in implicit stereotyping arising from different motivations to respond without prejudice. Amodio et al. (2008) found that, compared to other individuals, high IMS–low EMS individuals showed heightened conflict detection following stereotypical errors on the Weapons Identification Task (Payne, 2001). Given that knowing the correct response is a prerequisite for detecting conflict between correct and incorrect responses, these findings suggest that high IMS–low EMS individuals may be particularly successful at monitoring their responses for accuracy. If this is the case, then these same individuals should demonstrate higher estimates of the Quad model’s D parameter than other participants. On the other hand, high IMS–low EMS individuals may have weaker biased associations to begin with. Finally, different motivational styles may be related to differences in the ability to overcome associations. Just as enhanced detection would produce less implicit racial bias among high IMS–low EMS individuals, so too would reduced activation of biased associations or a greater ability to regulate those associations.

Method

Participants were 45 female students whose IMS and EMS scores fell within the top and bottom third of the distribution from an earlier mass testing session.[1] Plant and Devine’s (1998) IMS and EMS scales were used to assess internal and external motivation to respond without prejudice. As reported in Amodio et al. (2008), 12 participants had high IMS (M = 6.01, SD = 1.62) and low EMS (M = 4.76, SD = 2.46) scores, 17 participants were high in both IMS (M = 8.89, SD = 0.16) and EMS (M = 7.01, SD = 0.93), and 16

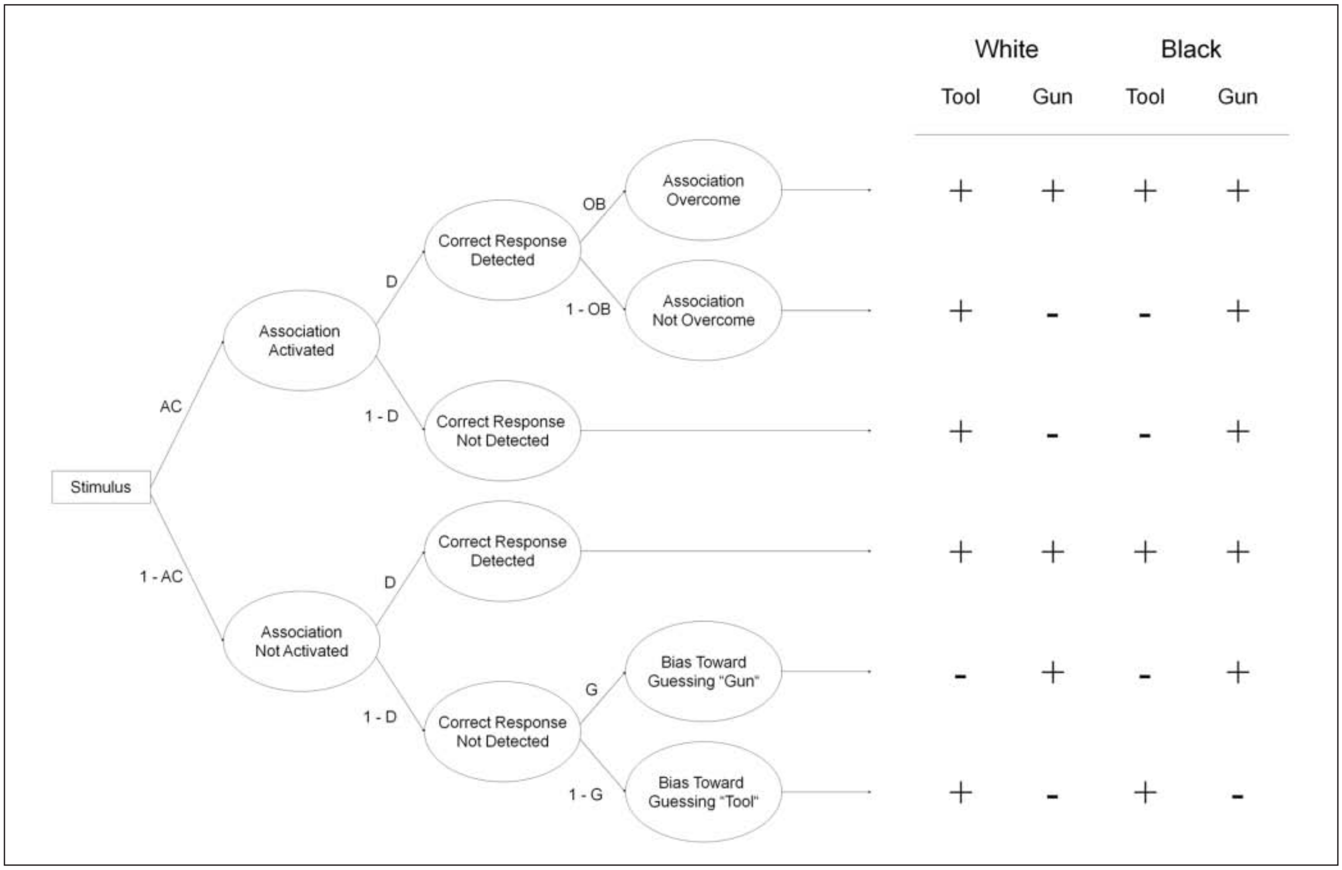

Figure 1. The Quadruple Process model (Quad model), as applied to Payne's (2001) Weapon Identification Task
Each path represents a likelihood of making a correct response. Parameters with lines leading to them are conditional upon all preceding parameters. The table on the right panel of the figure depicts correct (+) and incorrect (–) responses as a function of face prime (White vs. Black) and target stimulus (Tool vs. Gun).

participants had low IMS ($M = 2.70$, $SD = 1.06$) scores, collapsed across EMS scores ($M = 8.85$, $SD = 0.17$).[2]

Participants completed the Weapons Identification Task (Payne, 2001) in which the primes were two Black and two White male faces and target stimuli were four handguns and four hand tools. Each trial started with the presentation of a pattern mask for 1,000 ms, followed by the prime (a Black or White face) for 200 ms, the target (a gun or tool) for 200 ms, and finally a second mask for up to 2,000 ms from target onset. Participants were instructed to indicate, as quickly as possible, whether the target item was a gun or a tool by pressing one of two labeled buttons. If participants failed to respond within 500 ms of target onset, a message encouraging them to respond more quickly appeared. Participants completed 26 practice trials followed by 288 test trials. EEG was recorded throughout the task using procedures described by Amodio et al. (2008), but these measures were not included in the present analysis.

Results

A one-way ANOVA on stereotypical errors (i.e., misidentifying a tool as a gun following a Black face) on the Weapons Identification Task revealed a significant effect of motivation group, $F(2, 42) = 12.17$, $p < .001$.[3] High IMS–low EMS participants made a smaller percentage of errors ($M = 21.30$, $SD = 9.65$) than low IMS participants ($M = 44.36$, $SD = 47.20$), $t(42) = 4.52$, $p < .01$. Although the difference in stereotypical errors between high IMS–low EMS and high IMS–high EMS participants ($M = 26.39$, $SD = 11.25$) was in the expected direction, it did not reach significance, $t(42) = 1.01$, $p = .32$. This nonsignificant difference between high IMS–low EMS and high IMS–high EMS participants in stereotypical errors may mask important differences between these individuals in the processes underlying the expression of implicit bias. Such differences may be revealed via application of the Quad model.

The structure of the Quad model is depicted as a processing tree in Figure 1. In the tree, each path represents a likelihood. Processing parameters with lines leading to them are conditional on all preceding parameters. For instance, OB is conditional on both AC and D. The conditional relationships described by the model form a system of equations that predict the numbers of correct and incorrect responses in different conditions (e.g., compatible and incompatible trials). For example, the overall likelihood of producing an incorrect response to a tool preceded by a Black face prime is the sum of the three conditional probabilities: $[AC \times D \times (1 - OB)] +$

Table 1. Parameter Estimates for Weapons Task, Study 1

Parameter		High IMS–low EMS	High IMS–high EMS	Low IMS
		Estimate		
AC	Black–gun	.06	.00	.20
	White–tool	.22	.20	.09
D		.65	.52	.33
OB			.76	
G			.55	

IMS = Internal Motivation Scale; EMS = External Motivation Scale; AC = Activation of Associations; D = Detection; OB = Overcoming Bias: G = Guessing. The AC and D parameters were estimated within each motivation group, whereas the OB parameter and the G parameter were estimated across the three motivation groups. Goodness of model fit: $\chi^2(1) = 0.07$, $p = .79$.

[AC × (1 – D)] + [(1 – AC) × (1 – D) × G]. This equation sums the three possible paths by which an error can be returned in this case. The first part of the equation, AC × D × (1 – OB), is the likelihood that stereotypic associations are activated (AC) *and* detection succeeds (D) *and* OB fails (1 – OB). The second part of the equation, AC × (1 – D), is the likelihood that the stereotypical associations are activated (AC) *and* detection fails (1 – D). Finally, (1 – AC) × (1 – D) × G, is the likelihood that stereotypical associations are not activated (1 – AC) *and* detection fails (1 – D), *and* a bias toward guessing "gun" (G) produces an incorrect response. The respective equations for each item category (e.g., a Black person with a gun; a White person with a tool) are then used to predict the observed proportions of errors in a given data set. The model's predictions are then compared to the actual data to determine the model's ability to account for the data. A χ^2 estimate is computed for the difference between the predicted and observed errors. To best approximate the model to the data, the parameter values are changed through maximum likelihood estimation until they produce a minimum possible value of the χ^2. The final parameter values that result from this process are interpreted as relative levels of the processes.

With only 12 uniquely predicted categories of observations (Black and White primes paired with gun and tool targets, separated by motivation group), we were not able to generate separate estimates for all of the parameters for each motivation group within a single model. Thus, we tested two different versions of the Quad model.

In the first model, we estimated two AC parameters (Black–gun and White–tool AC) and one D parameter within each motivation group (high IMS–low EMS participants, high IMS–high EMS participants, and low IMS participants). Only one OB parameter and one G parameter were estimated, collapsed across all three groups, with G coded to represent a bias toward guessing "tool." This model enabled us to test for differences in AC and D across the three groups. This model fit the data, $\chi^2(1) = 0.07$, $p = .79$, with an overall error rate of 25.77%. As Table 1 shows, estimates for Black–gun AC were higher for low IMS participants than for high IMS–high EMS, $\Delta\chi^2(1) = 6.51$, $p = .01$, and high IMS–low EMS, $\Delta\chi^2(1) = 19.29$, $p < .0001$, participants. Unexpectedly, estimates for White–tool AC were lower for low IMS participants than for high IMS–high EMS, $\Delta\chi^2(1) = 5.14$, $p = .02$, and high IMS–low EMS, $\Delta\chi^2(1) = 7.28$, $p < .01$, participants. There were no differences in Black–gun or White–tool AC between the high IMS–low EMS participants and the high IMS–high EMS participants, $\Delta\chi^2(1) < 1.20$, ps > .27. However, these two groups differed with respect to the D parameter, with high IMS–low EMS participants showing greater ability to determine the correct response than high IMS–high EMS participants, $\Delta\chi^2(1) = 43.02$, $p < .00001$. The D estimate for the low IMS participants was significantly lower than for the high IMS–low EMS participants, $\Delta\chi^2(1) = 190.20$, $p < .00001$, and the high IMS–high EMS participants, $\Delta\chi^2(1) = 77.84$, $p < .00001$. These findings provide preliminary support that high IMS–low EMS participants are skilled in monitoring their responses for accuracy.

In the second model, we estimated a separate OB parameter for each motivational group, but only one D parameter collapsed across all groups. This allowed us to test for differences in OB among the groups. As in the first model, two AC parameters were estimated for each motivational group. This model did not fit the data, $\chi^2(1) = 165.35$, $p < .00001$. Thus, whereas the model with separate D parameters for each group was able to explain the data, the model with separate OB parameters was not.[4]

Discussion

In Study 1, we examined the processes underlying motivation-based differences in implicit bias on the Weapon Identification Task. Analysis of Amodio et al.'s (2008) data with the Quad model indicated that estimates of D were higher among high IMS–low EMS participants compared to high IMS–high EMS participants, who, in turn, showed higher estimates of D than low IMS participants. We also found that high IMS–low EMS and high IMS–high EMS participants had less biased associations than did those with low IMS.

The findings regarding the D parameter are consistent with Amodio et al.'s (2008) conclusion that high IMS–low EMS participants may be effective in controlling race bias because they have a more finely tuned conflict monitoring system. For conflict monitoring to occur, the correct and incorrect responses must be so identified. Together, these findings provide converging evidence that response monitoring is an important part of responding without bias. Detecting between appropriate and inappropriate responses may facilitate control over race bias as a response unfolds (Amodio et al., 2008), as well as in future situations (Monteith et al., 2002). A major advance of the current study is that it provided estimates of the response monitoring process that were independent of activation of automatic associations and inhibition. Moreover, estimates of all three processes were derived from a single behavioral measure. The current findings indicate that response monitoring processes contribute to motivation-based differences in implicit bias even when automatic associations and inhibition are taken into account.[5]

The findings regarding the AC parameter provide support for the role of automatic associations in motivation-based individual differences in implicit bias (e.g., Amodio et al., 2003; Devine et al., 2002; Glaser & Knowles, 2008; Johns et al., 2008; Moskowitz et al., 1999; Moskowitz et al., 2000; Park et al., 2008). These data suggest that individuals with varying beliefs and motivations may not possess similar implicit stereotypic associations, as posited in previous research (Amodio et al., 2008; Devine, 1989). In contrast to the present findings, the PD analyses reported by Amodio et al. (2008; also see Schlauch et al., 2009) showed only a nonsignificant trend for stronger automatic stereotyping bias among low IMS participants compared with either group of high IMS subjects. Our results indicated that high IMS–low EMS participants and high IMS–high EMS participants exhibited less activation of Black–gun associations compared to low IMS participants, a pattern of stereotype associations similar to previous observations of implicit racial attitudes and affective associations (Amodio et al., 2003; Devine et al., 2002). A detailed discussion of possible reasons for different PD and Quad model results is presented in the General Discussion. Surprisingly, low IMS participants showed less activation of White–tool associations than the other groups. However, there is no reason to expect White–tool associations to be particularly meaningful predictors in the context of motivations to control prejudice. Consistent with this idea, motivation group differences in the ERN in Amodio et al. (2008) were observed only for high-conflict trials involving Black primes (Black–tool trials) and not trials involving White primes.

Study 1 also provided novel insight into inhibition processes via the OB parameter. The current findings provide no support for the idea that differences in the ability to overcome the activated associations account for motivation-based individual differences in implicit bias. In fact, the model that specified separate OB parameters for the three groups was not able to account for the data. Nevertheless, our confidence in this conclusion would be strengthened if we were able to replicate our findings using a task design that enabled separate estimates of D and OB within the same model. This was an aim of Study 2.

Study 2

The goal of this study was to replicate the findings of Study 1 using a different measure of implicit associations, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). The advantage of using the IAT is that it provides enough degrees of freedom from unique trial types to predict separate estimates for each parameter within the same model. Another goal of Study 2 was to replicate the findings of Study 1 with a measure of implicit prejudice rather than stereotypes. Following Study 1, we hypothesized that high IMS–low EMS participants would show better D than high IMS–high EMS participants and low IMS participants. Based on the results of Study 1, we also predicted differences in the AC parameter across the three groups. We did not expect to find differences in OB or G as a function of motivation.

Method

A total of 72 undergraduates (41 females, *M* age = 20.20, *SD* = 1.52) whose IMS and EMS scores fell within the top and bottom third of the distribution from an earlier mass testing session received course credit for participating in the study. There were 23 high IMS–low EMS participants (IMS: *M* = 8.46, *SD* = 0.43; EMS: *M* = 2.87, *SD* = 0.75), 24 high IMS–high EMS participants (IMS: *M* = 8.62, *SD* = 0.38; EMS: *M* = 6.60, *SD* = 0.90), and 25 low IMS participants (IMS: *M* = 4.82, *SD* = 0.59; EMS: *M* = 5.10, *SD* = 1.61).

Participants completed a race evaluation IAT. Stimuli for the IAT were 10 images of faces (5 Black, 5 White) and 16 words (8 pleasant, 8 unpleasant). Participants first completed two 20-trial practice blocks, in which they discriminated pleasant from unpleasant words and Black from White faces. The third block was the "compatible" block, which consisted of 60 trials. Participants were instructed to press the right-hand (*i*) key whenever they saw a picture of a White person or a pleasant word and to press the left-hand (*e*) key whenever they saw a picture of a Black person or an unpleasant word. The keys used to categorize Black and White faces were switched in the remaining blocks. That is, participants categorized Black faces using the right-hand key and White faces using the left-hand key. The fourth block was a practice block in which participants discriminated Black from White faces. The final block consisted of 60 "incompatible" trials, in which "Black" shared a response key with the evaluative dimension "pleasant." Participants who respond more quickly when "Black" shares a key with "unpleasant" ("compatible" trials) than when it shares a key with "pleasant" ("incompatible" trials) are thought to have an implicit preference for Whites relative to Blacks (Greenwald et al., 1998). Category labels remained on the top left and right of the screen throughout the task, whereas stimulus pictures and words appeared in the center of the screen. A red *X* appeared whenever participants made an error, and they were required to correct it before moving on to the next trial.

Results

IAT scores were calculated according to the algorithm described by Greenwald, Nosek, and Banaji (2003). Trial latencies greater than 10,000 ms were dropped from analysis prior to computing separate mean latencies for the compatible

Table 2. Parameter Estimates for Black–White Implicit Association Test, Study 2

Parameter		Estimate		
		High IMS–low EMS	High IMS–high EMS	Low IMS
AC	Black–unpleasant	.05	.09	.11
	White–pleasant	.07	.11	.15
D		.93	.87	.86
OB		.80	.87	.71
G		.55	.51	.51

IMS = Internal Motivation Scale; EMS = External Motivation Scale; AC = Activation of Associations; D = Detection; OB = Overcoming Bias: G = Guessing. Goodness of model fit: $\chi^2(3) = 3.92$, $p = .27$.

and incompatible blocks. Because the IAT contained a built-in error penalty, no further penalty was applied to error latencies (Greenwald et al., 2003). The difference between the mean compatible and incompatible latencies was then divided by the pooled standard deviation of all critical trials to produce IAT scores, such that higher scores indicate a stronger preference for Whites relative to Blacks. A one-way ANOVA revealed a significant effect of regulation group, $F(2, 69) = 4.50$, $p = .02$. High IMS–low EMS participants showed less IAT bias ($M = 0.47$, $SD = 0.42$) than high IMS–high EMS participants ($M = 0.69$, $SD = 0.38$), $t(69) = 2.12$, $p < .04$ and low IMS participants ($M = 0.77$, $SD = 0.28$), $t(69) = 2.90$, $p < .01$. The mean IAT scores of high IMS–high EMS participants and low IMS participants were not significantly different, $t(69) = 0.73$, $p = .47$.

For each group, we calculated parameter estimates of AC, D, OB, and G. The G parameter was coded so that higher scores represented a bias toward guessing with the positive ("pleasant") key. Two separate AC parameters were estimated: one measuring the extent to which associations between "Black" and "unpleasant" were activated in performing the task and another measuring the extent to which associations between "White" and "pleasant" were activated. The Quad model fit the data, $\chi^2(3) = 3.92$, $p = .27$, with an overall error rate of 8.07%.[6]

Replicating the results of Study 1, high IMS–low EMS participants showed better D than high IMS–high EMS participants, $\Delta\chi^2(1) = 12.77$, $p < .001$, and low IMS participants, $\Delta\chi^2(1) = 17.59$, $p < .0001$ (see Table 2). High IMS–high EMS participants and low IMS participants did not differ in D, $\Delta\chi^2(1) = 0.29$, $p = .60$. In addition, compared to the low IMS participants, estimates for Black–unpleasant, $\Delta\chi^2(1) = 5.58$, $p = .02$, and White–pleasant association activation, $\Delta\chi^2(1) = 8.99$, $p < .01$, were significantly lower for the high IMS–low EMS participants. Levels of Black–unpleasant AC and White–pleasant AC also were lower for high IMS–low EMS participants than for high IMS–high EMS participants, although these differences were marginal, $\Delta\chi^2(1) = 2.97$, $p = .08$, or not significant, $\Delta\chi^2(1) = 2.40$, $p = .12$, respectively. No other AC differences were significant, $\Delta\chi^2(1) < 0.40$, $ps > .50$. There were no differences in OB or G among the three groups, $\Delta\chi^2(1) < 0.70$, $ps > .40$.[7]

Discussion

The Study 2 results replicated the findings that high IMS–low EMS participants were more likely to detect appropriate and inappropriate responses than high IMS–high EMS participants and low IMS participants. In addition, high IMS–low EMS participants showed lower levels of association activation than low IMS participants. In contrast to Study 1, there was some evidence that high IMS–low EMS participants also had lower AC estimates than high IMS–high EMS participants (see especially Note 6). It is possible that high IMS–low EMS participants and high IMS–high EMS participants differ in activation of evaluative associations (as assessed by the IAT) but not stereotypic associations (e.g., as assessed by the Weapons Identification Task; see Amodio et al., 2008, for a discussion of this distinction). This is a question for future research. In any case, the results of both studies indicate that levels of activated associations differ among individuals depending on their motivations to respond without prejudice.

The present study extended Study 1 by producing separate estimates for the D and OB parameters for each group within the same statistical model. There was no evidence of OB differences among the three groups. The results of Study 1 suggested that OB did not account for differences among the three groups. The current study similarly found no evidence for such differences using a more direct test. The current findings thus strengthen our conclusion that observations of motivation-based differences in implicit bias are a function of D and AC.

General Discussion

Previous research has reported that individuals who are high in chronic egalitarian goals, are high in motivation to control prejudice, or are internally (but not externally) motivated to respond without prejudice show lower levels of implicit racial bias than other individuals (e.g., Devine et al., 2002; Dunton & Fazio, 1997; Johns et al., 2008; Moskowitz et al., 1999; Moskowitz et al., 2000). A variety of theoretical accounts have been proposed to explain these effects. The purpose of the present research was to directly test these accounts. Specifically, we tested whether those with high internal but not external motivations to respond without prejudice have lower levels of activated associations, greater inhibition of automatic associations, or enhanced response monitoring compared to other individuals. We distinguished among these accounts by applying the Quad model, which separately measures and disentangles the contributions of these processes to performance on implicit measures. Across two studies, high IMS–low EMS individuals showed enhanced detection of appropriate responses compared to other individuals. They also showed less activation of biased associations than low IMS participants in both studies and than high IMS–high EMS participants in Study 2, in which evaluative bias was examined. Neither study showed group differences in overcoming biased associations.

These findings expand prior research highlighting the important role of motivation-based differences in response monitoring. Previous results showed that high IMS–low EMS individuals are particularly able to detect when there are conflicting responses (Amodio et al., 2008), thereby increasing the likelihood of responding appropriately. The present results show that these same individuals are more successful at determining appropriate responses (e.g., Amodio et al., 2004; Monteith et al., 2002), an aspect of response monitoring that is critical for detecting conflict between competing responses.

Important questions for future research concern the causal relationships among motivations to respond without bias, response monitoring, and the activation of biased associations. It is not clear if high IMS–low EMS individuals have less biased associations because of extensive practice at monitoring and inhibiting biased responses or if they simply never had biased associations in the first place. It also is possible that having less biased associations facilitates the development of effective response monitoring and that this could produce the D parameter results. Another possibility is that a third variable (e.g., egalitarian beliefs) could lead to both less biased associations and more effective response monitoring. Finally, another important question for future research is the extent to which successful response monitoring is a reflection of motivation, skill, or both. The analysis of flanker task data (see Note 4) suggests that it is not a general cognitive ability. Still, it is possible that high D represents motivation, ability, or both within the domain of intergroup bias.

The present results failed to provide support for the hypothesis that differences in task performance between motivation groups were the result of differences in their ability to overcome biased associations. Had high IMS–low EMS participants been effectively overriding automatic associations, then we would have expected to see greater OB among these participants. Rather, the findings suggest that internally motivated individuals are simply less likely to have biased associations activated in the first place. Whether these lower levels of association activation are the result of extended prior practice at regulating biased responses (e.g., Monteith et al., 2002) is a question for future research.

Assuming this null result is meaningful, we can offer only a speculative account for why high IMS–low EMS individuals would not show enhanced OB. It may be that, because they have weaker associations in the first place, high IMS–low EMS individuals simply have fewer opportunities to practice overcoming bias. This could counteract any inherent motivational advantage they might have. The motivational profile alone may be sufficient to increase the likelihood of determining appropriate behavior and of noticing the (relatively infrequent) response conflicts experienced by these individuals. However, it may be that direct practice is required to enhance inhibition. Further research may shed light on this speculation.[8]

Application of the Quad model also revealed different results than applications of the PD procedure (Amodio et al., 2008; Schlauch et al., 2009). Although application of PD has failed to show motivation-based differences in an automatic stereotyping component of bias, the present results did reveal some differences. What might account for the discrepant findings when applying PD versus the Quad model? One possibility lies in the fact that the automatic processes estimated by PD and the Quad model differ in an important way. Specifically, the automatic bias estimated by the PD model represents a highly constrained form of automaticity. According to PD, when the automatic bias (A) and the controlled detection response (C) are in conflict, the automatic bias may influence responses only when participants have failed to detect the correct response. That is, if the respondent is able to accurately discern whether the stimulus object is a gun or a tool (e.g., in the Weapons Identification Task), then the automatic bias will have no influence on the response. Thus, PD results refer to a specific type of automatic process that is subordinate to controlled processes (for a review, see Sherman, Klauer, & Allen, 2010).

In contrast, the Quad model does not constrain automatic bias in this way. According to the Quad model, whether competing automatic associations (AC) or controlled response detection processes (D) determine behavior depends on the success of the overcoming bias (OB) process. If OB succeeds, then D determines behavior, regardless of automatic bias; if OB fails, then AC determines behavior, regardless of controlled response detection. Thus, the AC process measured by the Quad model may be either superordinate or subordinate to D. It may be that differently motivated individuals vary in the more robust type of automaticity measured by the Quad model but not in the subordinate automaticity estimated by PD.

Of course, another important difference between the Quad model and PD is the specification of the OB parameter in the Quad model. OB plays a critical role in the Quad model, acting as the arbiter of whether AC or D determines responses when they are in conflict. Specification of the OB (and G) parameter necessarily produces an estimate of automatic processing that differs from that estimated by PD. This also may contribute to the different results demonstrated by the two procedures. We do not take these results to mean that the AC parameter of the Quad model is in any sense superior to the A parameter of PD. The parameters of the models measure distinct constructs, both of which have been shown to be important components of implicit bias.

Conclusion

Understanding why some people show less bias on implicit measures than others has preoccupied researchers concerned with prejudice reduction. Different explanations have been advanced for findings that implicit bias is a function of egalitarian goals and motivations to respond without prejudice. The present research indicates that low levels of implicit bias are associated with weak activation of racial associations as well as success at monitoring responses for accuracy. Specifying the key process differences among people offers valuable insights into the means by which some individuals are able to avoid implicit bias and suggests how interventions might work to decrease bias.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a grant from the National Science Foundation (BCS 0820855) to the second author.

Notes

1. This study was a reanalysis of Amodio, Devine, and Harmon-Jones (2008, Study 1), which measured physiological responses and selected females to control for variability associated with sex.
2. Among those low in internal motivation to respond without prejudice (IMS), level of external motivation to respond without prejudice (EMS) does not influence the extent of implicit bias (e.g., Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002). As such, level of EMS was collapsed among participants low in IMS. In contrast, among those high in IMS, level of EMS does influence the extent of implicit bias. The main purpose of the present research was to better understand the reasons for this effect.
3. This difference in stereotypical errors was not reported by Amodio et al. (2008).
4. Although our goal was not to compare the models, we calculated fit indices for the Quad and Process Dissociation (PD) models for interested readers. Specifically, we generated estimates of Akaike information criterion (AIC) and Bayes information criterion (BIC), two metrics of model fit that correct for model complexity (for a review, see Myung, 2000). The data from Study 1 fit both the Quad model, $\chi^2(1) = 0.07$, $p = .79$, and the PD model, $\chi^2(3) = 1.78$, $p = .62$. Estimates of AIC and BIC were very similar for the two models (AIC = 14553.78 for the Quad model vs. 14551.49 for the PD model, BIC = 14635.95 for the Quad model vs. 14618.71 for the PD model). The data from Study 2, which measured implicit bias with an Implicit Association Test (IAT), fit the Quad model, $\chi^2(3) = 3.97$, $p = .27$, but not the PD model, $\chi^2(3) = 55.97$, $p < .001$. Nevertheless, estimates of AIC and BIC were again very similar (AIC = 4648.31 for the Quad model vs. 4688.31 for the PD model, BIC = 4754.27 for the Quad model vs. 4751.89 for the PD model). These data corroborate previous work showing that the PD model provides slightly better fit on priming measures, whereas the Quad model provides better fit for IAT performance, though these differences in model fit are small (Sherman et al., 2008). We believe that both models can be used to answer important (but different) questions about the processes contributing to implicit bias, a point to which we return in the General Discussion.
5. Amodio et al. (2008, Study 2) gave participants a flanker task that included congruent (<<<<< or >>>>>) and incongruent (<<><< or >><>>) trials. From these trials, Amodio et al. computed PD estimates of automatic and controlled processing in performing the flanker task. Results showed that high IMS–low EMS and high EMS–high EMS participants did not differ on controlled processing (C). In contrast, both of those groups showed more control than low IMS participants. We analyzed these flanker data with the Quad model and replicated these results. Specifically, we found that high IMS–low EMS and high IMS–high EMS participants did not differ on the D parameter for the flanker task. However, both groups showed higher D than low IMS participants. Thus, whereas high IMS–low EMS participants had higher D than high IMS–high EMS participants on the measure of implicit bias, they did not differ on the flanker task, suggesting that the D advantage is specific to the domain of implicit stereotyping and does not represent a more general cognitive ability.
6. A χ^2 analysis showed that the errors did not differ between the target and attribute categories in the compatible trials, χ^2s$(1) < 1.10$, *p*s > .31. As such, for modeling these trials, errors for White and pleasant items were collapsed, as were errors for Black and unpleasant items.
7. When each individual participant performs relatively few trials (as in the current case), parameter estimates derived from aggregated data are more accurate than parameter estimates derived from each participant separately (e.g., Cohen, Sanborn, & Shiffrin, 2008). As such, our analyses utilized aggregated data. However, because they do not account for individual differences, one concern with aggregated data is that they may violate assumptions of interparticipant parameter homogeneity. To account for this threat, we applied Klauer's (2006) latent-class hierarchical multinomial modeling procedure to the data. This procedure computes parameter estimates separately for latent classes of participants, permitting examination of the influence of parameter heterogeneity on results. Application of this procedure to the Study 2 data reproduced the findings from the standard aggregate analysis, except that the differences in Black–unpleasant and White–pleasant associations between high IMS–low EMS and high IMS–high EMS participants reached significance, $t(69) > 2.57$, $p < .05$. These analyses indicate that the results were not an artifact of parameter heterogeneity. The latent-class hierarchical procedure could not be applied to the Study 1 data because of insufficient degrees of freedom in the model.
8. Given a conflict between biased associations and a detected correct response, Overcoming Bias (OB) measures the likelihood that the correct response tendency will override the biased response tendency. However, OB does not assess the potential role of inhibitory processes that would diminish the extent to which biased associations are activated in the first place. In the Quad model, the operation of such an inhibition process could not be distinguished from the absence of biased associations.

References

Allen, T. J., Sherman, J. W., & Klauer, K. C. (2010). Social context and the self-regulation of implicit bias. *Group Processes and Intergroup Relations*, *13*, 137-149.

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology*, *94*, 60-74.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology*, *84*, 738-753.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science*, *15*, 88-93.

Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). Alternative mechanisms for regulating racial responses according to internal vs. external cues. *Social Cognitive and Affective Neuroscience*, *1*, 26-36.

Bartholow, B. D., Dickter, C. L., & Sestir, M. A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology*, *90*, 272-287.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin and Review*, *6*, 57-86.

Beer, J. S., Stallen, M., Lombardo, M. V., Gonsalkorale, K., Cunningham, W. A., & Sherman, J. W. (2008). The Quadruple Process model approach to examining the neural underpinnings of prejudice. *Neuroimage*, *43*, 775-783.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624-652.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*, 747-749.

Cohen, A. L., Sanborn, A. N., & Shiffrin, R. M. (2008). Model evaluation using grouped or individual data. *Psychonomic Bulletin and Review*, *15*, 692-712.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality and Social Psychology*, *89*, 469-487.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5-18.

Devine, P. G., Plant, E., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, *82*, 835-848.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*, 316-326.

Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice. *Journal of Experimental Social Psychology*, *44*, 164-172.

Gonsalkorale, K., Allen, T. J., Sherman, J. W., & Klauer, K. C. (2010). Mechanisms of exemplar exposure effects on implicit attitudes. *Social Psychology*, *41*, 158-168.

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, *45*, 410-414.

Gonsalkorale, K., von Hippel, W., Sherman, J. W., & Klauer, K. C. (2009). Bias and regulation of bias in intergroup interactions: Implicit attitudes toward Muslims and interaction quality. *Journal of Experimental Social Psychology*, *45*, 161-166.

Gordijn, E. H., Hindriks, I., Koomen, W., Dijksterhuis, A., & van Knippenberg, A. (2004). Consequences of stereotype suppression and internal suppression motivation: A self-regulation approach. *Personality and Social Psychology Bulletin*, *30*, 212-224.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197-216.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513-541.

Johns, M., Cullum, J., Smith, T., & Freng, S. (2008). Internal motivation to respond without prejudice and automatic egalitarian goal activation. *Journal of Experimental Social Psychology*, *44*, 1514-1519.

Klauer, K. C. (2006). Hierarchical multinomial processing tree models: A latent-class approach. *Psychometrika*, *71*, 7-31.

Legault, L., Green-Demers, I., & Eadie, A. L. (2009). When internalization leads to automatization: The role of self-determination in automatic stereotype suppression and implicit prejudice regulation. *Motivation and Emotion*, *33*, 10-24.

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development and operation of cues for control. *Journal of Personality and Social Psychology*, *83*, 1029-1050.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*, 167-184.

Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Preconsciously controlling stereotyping: Implicitly activated egalitarian goals prevent the activation of stereotypes. *Social Cognition*, *18*, 151-177.

Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology*, *44*, 190-204.

Park, S. H., Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice moderates the effect of cognitive depletion on unintended discrimination. *Social Cognition*, *26*, 401-419.

Payne, B. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181-192.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811-832.

Schlauch, R. C., Lang, A. R., Plant, E. A., Christensen, R., & Donohue, K. F. (2009). Effect of alcohol on race-biased responding: The moderating role of internal and external motivations to respond without prejudice. *Journal of Studies on Alcohol and Drugs*, *70*, 328-336.

Sherman, J. W. (2006). On building a better process model: It's not only how many, but which ones and by which means. *Psychological Inquiry*, *17*, 173-184.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, *115*, 314-335.

Sherman, J. W., Klauer, K. C., & Allen, T. J. (2010). Mathematical modeling of implicit social cognition: The machine in the ghost. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement*, *theory*, *and applications* (pp. 156-175). New York, NY: Guilford.